DOCUMENT RESUME

ED 264 280 TM 850 766

AUTHOR Steele, D. Joyce
TITLE A Descriptive Comparison of Test Items Utilized in Pilot and Live Administrations of the Alabama High School Graduation Examination.
PUB DATE 8 Nov 85
NOTE 18p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (New Orleans, LA, November 8, 1985).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; *Comparative Testing; *Difficulty Level; Graduation Requirements; High Schools; Language Tests; Mathematics Tests; *Minimum Competency Testing; *Pretesting; Reading Tests; Sampling; State Programs; Test Bias; *Test Construction; Testing Programs; *Test Validity
IDENTIFIERS *Alabama High School Graduation Examination

ABSTRACT

This paper contains a comparison of descriptive information based on analyses of pilot and live administrations of the Alabama High School Graduation Examination (AHSGE). The test is composed of three subject tests: Reading, Mathematics, and Language. The study was intended to validate the test development procedure by comparing difficulty levels of common test items measuring competencies on the initial item pilot and on the live administrations. Data were collected during two phases of pilot testing and four live administrations. Public schools in Alabama with at least 50 students were randomly sampled to participate in the pilot procedure. Data were analyzed, and difficulty levels and discrimination indexes for all pilot items were calculated, based on the proportion of students completing the item correctly. A second pilot test involved testing of forms. Following live administrations, common items were compared with the initial item pilot. The comparison demonstrated the stability of the difficulty level and supported the validity of the test development procedure. Although a higher percentage of students passed items on the live administration, the slight increase in percentage seemed to be consistent from one form to another. (LMO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

A Descriptive Comparison of Test Items Utilized in Pilot and Live Administrations of the Alabama High School Graduation Examination

ED264280

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Steele

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

D. Joyce Steele, Ph.D.
Evaluation and Assessment Laboratory
The University of Alabama
P. O. Box Q/330 Bibb Graves Hall
University, Alabama 35486
(205) 348-7575

TM 850 766

Mid-South Educational Research Association
New Orleans, Louisiana
November 8, 1985

A Descriptive Comparison of Test Items
Utilized in Pilot and Live Administrations
of the Alabama High School Graduation Examination

Introduction

This paper contains a comparison of descriptive information based on analyses of the pilot and live administrations of The Alabama High School Graduation Examination (AHSGE). The AHSGE is a product of decisions made in 1977 by the Alabama State Board of Education; moreover, in 1981, new requirements for graduation from high school were approved which mandated that students pass a minimum competency examination in addition to earning the required Carnegie units (McLean, 1984). The AHSGE is composed of three subject tests: Reading, Mathematics, and Language. The pass score for each subject test of the AHSGE was established using empirical and judgmental methods (Halpin, 1984; Roth, 1984) based on data from the initial pilot administration; therefore, it is essential that the live administration data are consistent with the initial pilot data. The purpose of this study was to validate the test development procedure by comparing the difficulty level of common test items measuring competencies on the initial item pilot and on the live administration.

The data were collected during two phases of pilot testing and four live administrations of the AHSGE. The following section summarizes the item pilot procedure, the form pilot procedure, and the comparisons of pilot data to the live administration data.

Item Pilot Procedure

The initial item pilot procedure was conducted by the Educational Testing Service. According to an October 1, 1982 memorandum from the State

Superintendent of Education, Wayne Teague, the item pilot procedure was conducted November 16-18, 1982. All public schools in Alabama with at least 50 students were among the population which was randomly sampled to participate in the item pilot procedure. The state of Alabama was divided into four regions determined by student population. Fourteen schools and five alternates were randomly selected from each of the four regions.

There were ten booklets of test items. There were three test booklets for reading items, another three test booklets for mathematics items, and four test booklets for language.

According to the records of the Alabama State Department of Education, approximately one percent of the 55,000 eleventh grade Alabama students (for the 1982-83 academic year) were randomly selected. Table 1 shows the number of students involved in the item pilot at school, district, and state levels.

Table 1

Form	Number of students		
	School	District	State
A Reading, Mathematics, Language	10	140	560
B Reading, Mathematics, Language	10	140	560
C Reading, Mathematics, Language	10	140	560
D Language	10	140	560
Total	40	560	2240

Note. This table is based on information included with a memorandum from the Alabama State Superintendent of Education to City and County Superintendents of systems.

Analysis of Pilot Items

Once the pilot tests were administered, the Evaluation and Assessment Laboratory at The University of Alabama analyzed the data from results. The difficulty level and discrimination index for all pilot items were calculated. The difficulty level indicated the proportion of students who completed the item correctly. The discrimination index shows how well the item discriminates between students who score high and students who score low on all test items in the test booklets. The pilot items were analyzed to determine whether they were sexually or ethnically biased. All questionable items were eliminated from the item pool.

The next step in the pilot procedure involved the pilot testing of forms. The next section of this paper will provide a brief overview of the procedures involved.

Form Pilot Procedures

Two equivalent forms of the AHSGE were developed from the pool of pilot items for reading, mathematics, and language. The reading pilot forms were forms 11 and 12; the mathematics pilot forms were forms 23 and 24, and the language pilot forms were forms 35 and 36. The reading, mathematics, and language pilot forms were administered on April 26, 27, and 28, 1983, respectively. The same 1982-83 eleventh grade class was used for the form pilot; however, no school participated in both the item pilot and the form pilot.

Sampling Procedure

The approximately 450 schools were arranged into five achievement strata based upon the average California Achievement Tests performance of the 1981-82 tenth grade class (the eleventh grade class of 1982-83). Once the schools were arranged into five achievement strata, eight schools and two alternate schools were selected from each strata using random sampling procedures. A

proportional sample was taken from each of the eight schools until the sample size was approximately 240 per strata.

An alphabetized list of eleventh grade students from the randomly selected schools was obtained. The student lists were numbered and random numbers were generated for each form of the test. One set of students received Forms 11, 23, and 35 for reading, mathematics, and language respectively, and the other set of students in the school received the other forms for each subject area. Approximately 600 students received Forms 11, 23, and 35, and approximately 600 students received Forms 12, 24, and 36. There were about 1,200 students in the total sample from the approximately 55,000 student population for the 1982-83 eleventh grade class. Appendix B contains a list of randomly selected schools and alternate schools for each of the five achievement strata. Also, Appendix B provides the population and sample size for each randomly selected school. Although each of the five strata had two alternate schools available in case of an emergency, none of the alternate schools were used.

Analysis

Table 2 describes the summary statistics for the pilot forms. The reliability coefficients ranged from .95 to .96 for all pilot forms.

Table 2

Summary Statistics for the Pilot Forms of the AHSGE

Subject Form	n	Number of Items	M	SD	SE	Cronbach Alpha
Reading						
Form 11	609	80	72.03	10.66	.43	.95
Form 12	610	80	71.37	11.14	.45	.95
Mathematics						
Form 23	610	95	72.77	16.26	.66	.95
Form 24	610	95	73.18	16.11	.65	.96
Language						
Form 35	615	119	99.12	17.54	.71	.96
Form 36	616	119	98.32	18.58	.75	.96

Note. The sample included about 1231 students out of a population of about 55,000 eleventh graders in the 1982-83 academic year.

Live Administration of the AHSGE

In October 1983, over 48,000 eleventh grade Alabama students participated in the first live administration of the AHSGE. The student count for the live administrations for Spring 1984, Fall 1984 and Spring 1985 were approximately 10,000, 52,000 and 11,000 respectively. Since the Fall 1983 and Spring 1985 were essentially the same items as those used in the form pilot in April, 1983, the difficulty level from these two forms were used to compare the difficulty level with the difficulty level on the initial item pilot. The purpose of this comparison is to demonstrate the stability of the difficulty

level from the initial pilot to the live administration of the AHSGE. This demonstration of the stability of the difficulty level supports the validation of the test development procedure. This validation is essential to the credibility of the AHSGE since each form of the test was built with the assumption that the difficulty level of pilot items would be a close estimate of the actual difficulty level on a live administration. As can be observed from Tables 3 through 8, the difficulty level remains stable from the initial pilot administration to the live administration. Only students attempting the test for the first time were included in this data. As could be expected, a slightly higher percentage of students are getting the items correct on the live administration than on the initial item pilot. Students taking the initial item pilot were not required to pass the AHSGE in order to get a high school diploma; whereas, students in the live administration knew that their high school diploma was at stake if they did not pass the AHSGE. Also, the students in the live administrations had known since the ninth grade that passing the AHSGE was going to be a requirement for the high school diploma.

Although a higher percentage of students are passing items on the live administration than on the initial item pilot, the difficulty level appears to be stable and the slight increase in the percentage of students passing the items seems to be consistent from one form to another form.

Table 3
A Comparison of the Average Difficulty Level (percentage correct) on the Initial Pilot and the Live Administration of the AHSGE for Reading

Competency Number	Competency Description	# Items	Form 11 Item Pilot	Fall 83 Live Administration
I-1,2; IV-13	WORDS IN DAILY LIFE; FILLING OUT FORMS	8	88.03	93.00
II-1	CONTEXT CLUES	5	83.98	88.60
III-1	PREFIXES & SUFFIXES	4	87.95	93.00
III-2	ABBREVIATIONS	4	93.73	96.25
IV-1	ALPHABETICAL ORGANIZATION	6	86.52	91.83
IV-3	REFERENCE SOURCES	4	87.35	93.50
IV-4	BOOK SECTIONS	4	81.28	86.74
IV-5	NEWSPAPER SECTIONS	4	92.78	95.25
IV-6	MAIN IDEA: PASSAGES	5	83.52	90.60
IV-7	FIND DETAILS IN A WRITTEN PASSAGE	5	87.90	93.00
IV-8	USE MAPS, GRAPHS, CHARTS & TABLES	6	91.05	94.83
IV-9, 10, 11	MAIN IDEA & DETAIL: GOV, LAW, MON MGT, NUTR	9	86.69	92.89
IV-14	ADS, LABELS, CATALOGS AND DIRECTORIES	6	91.20	94.50
V-1	FOLLOWING DIRECTIONS & INSTRUCTIONS	6	89.45	93.83
VI-1	FACT & OPINION	4	79.85	87.50
		Total = 80	87.53	92.50

Table 4
A Comparison of the Average Difficulty Levels (percentage correct) on the Initial Pilot and the Live Administration of the AHSGE for Reading

Competency Number	Competency Description	# Items	Form 12 Item Pilot	Spring 85 Live Administration
I-1,2; IV-13	WORDS IN DAILY LIFE: FILLING OUT FORMS	8	88.81	92.00
II-1	CONTEXT CLUES	5	83.18	90.20
III-1	PREFIXES & SUFFIXES	4	87.15	88.25
III-2	ABBREVIATIONS	4	93.13	95.75
IV-1	ALPHABETICAL ORGANIZATION	6	86.48	91.67
IV-3	REFERENCE SOURCES	4	87.35	93.00
IV-4	BOOK SECTIONS	4	81.00	83.00
IV-5	NEWSPAPER SECTIONS	4	92.65	96.00
IV-6	MAIN IDEA: PASSAGES	5	81.58	86.20
IV-7	FIND DETAILS IN A WRITTEN PASSAGE	5	91.90	95.20
IV-8	USE MAPS, GRAPHS, CHARTS & TABLES	6	90.95	94.00
IV-9, 10, 11	MAIN IDEA & DETAIL: GOV, LAW, MON MGT, NUTR	9	86.04	93.33
IV-14	ADS, LABELS, CATALOGS AND DIRECTORIES	6	91.12	95.50
V-1	FOLLOWING DIRECTIONS & INSTRUCTIONS	6	88.43	93.67
VI-1	FACT & OPINION	4	82.40	88.25
		Total = 80	87.56	92.00

Table 5

A Comparison of the Average Difficulty Levels (percentage correct) on the Initial Pilot and the Live Administration of the AHSGE for Mathematics

Competency Number	Competency Description	# Items	Form 23 Item Pilot	Fall 83 Live Administration
I-1	READ & WRITE: MONEY VALUES, NUMBERS	3	92.47	95.33
I-2	COMPARE DECIMALS: COMPARE FRACTIONS	3	66.07	74.67
I-3	COMPUTATION: WHOLE NUMBERS	6	92.72	95.83
I-4	COMPUTATION: DECIMALS	6	84.68	90.17
I-5	COMPUTATION: FRACTIONS	6	60.88	72.17
I-6	CONVERSION: DECIMALS AND PERCENTS	3	60.57	78.33
I-7	APPLICATION: RATIO, %, DISC, TAXES, COMM	4	57.00	67.25
I-8	COMVERSION: FRACTIONS & DECIMALS	4	62.80	72.25
I-9	AVERAGES OF WHOLE & DECIMAL NUMBERS	3	75.93	85.67
I-10	CONVERSION: UNITS OF MEASURE	4	43.90	55.00
II-1	SOLVE PROBLEMS INVOLVING TIME	3	66.60	75.67
II-2	MEASURE LENGTHS BY USE OF A RULER	4	71.45	85.75
II-3	CAPACITIES	3	77.00	86.00
II-4	FIND PERIMETERS	4	89.35	92.50
II-5	FIND RECTANGULAR AREAS	4	31.50	50.25
III-1,2	INTERPRET: GRAPHS, TABLES, CHARTS	4	78.58	85.50

III-3	INTERPRET: SCALE DRAWINGS	3	78.07	86.33
IV-1	APPROXIMATIONS BY ROUNDING NUMBERS	3	82.13	90.67
V-1,2,3	GEOMETRIC FORMS AND CONCEPTS	6	89.98	87.33
VI-1,2,3	DETRM CHAG: COMP COST: COMP CASH/CRED PRCH	6	72.83	86.00
VI-4	ALLOCATE TIME AND MONEY	3	63.77	74.33
VI-5	COMPLETE CHECKS AND MONEY ORDERS	3	76.03	90.00
VI-6	SOLVE RATE, TIME, DISTANCE PROBLEMS	3	76.50	89.00
VI-7	COMPUTE SALARIES AND WAGES	4	73.43	87.25
		Total = 95	72.47	81.58

Table 6

A Comparison of the Average Difficulty Levels (percentage correct) on the Initial Pilot and the Live Administration of the AHSGE for Mathematics

Competency Number	Competency Description	# Items	Form 24 Item Pilot	Spring 85 Live Administration
I-1	READ & WRITE: MONEY VALUES, NUMBERS	3	93.40	95.67
I-2	COMPARE DECIMALS: COMPARE FRACTIONS	3	65.93	76.00
I-3	COMPUTATION: WHOLE NUMBERS	6	92.93	95.67
I-4	COMPUTATION: DECIMALS	6	85.10	92.50
I-5	COMPUTATION: FRACTIONS	6	60.90	66.83
I-6	CONVERSION: DECIMALS AND PERCENTS	3	58.77	75.33
I-7	APPLICATION: RATIO, %, DISC, TAXES, COMM	4	58.85	68.75
I-8	COMVERSION: FRACTIONS & DECIMALS	4	62.80	73.75
I-9	AVERAGES OF WHOLE & DECIMAL NUMBERS	3	76.10	83.67
I-10	CONVERSION: UNITS OF MEASURE	4	45.20	49.50
II-1	SOLVE PROBLEMS INVOLVING TIME	3	70.33	78.67
II-2	MEASURE LENGTHS BY USE OF A RULER	4	69.83	80.00
II-3	CAPACITIES	3	77.50	92.00
II-4	FIND PERIMETERS	4	89.10	91.75
II-5	FIND RECTANGULAR AREAS	4	29.28	48.00
III-1,2	INTERPRET: GRAPHS, TABLES, CHARTS	4	78.48	83.50

III-3	INTERPRET: SCALE DRAWINGS	3	80.20	83.33
IV-1	APPROXIMATIONS BY ROUNDING NUMBERS	3	82.50	88.00
V-1,2,3	GEOMETRIC FORMS AND CONCEPTS	6	89.42	94.83
VI-1,2,3	DETRM CHAG: COMP COST: COMP CASH/CRED PRCH	6	73.10	84.00
VI-4	ALLOCATE TIME AND MONEY	3	63.60	73.33
VI-5	COMPLETE CHECKS AND MONEY ORDERS	3	72.77	87.33
VI-6	SOLVE RATE, TIME, DISTANCE PROBLEMS	3	76.30	88.67
VI-7	COMPUTE SALARIES AND WAGES	4	73.08	84.50
		Total = 95	72.48	80.92

Table 7

A Comparison of the Average Difficulty Levels (percentage correct) on on the Initial Pilot and the Live Administration of the AHSGE for Language

Competency Number	Competency Description	# Items	Form 35 Item Pilot	Fall 83 Live Administration
I-1	PRONOUN-ANTECEDENT AGREEMENT	9	93.38	95.89
I-2	CORRECT FORMS OF NOUNS AND VERBS	15	79.94	82.93
II-1	IDEAS IN SEQUENTIAL ORDER	5	80.94	85.60
II-2	SENTENCES ORGANIZED INTO PARAGRAPH	9	80.49	84.56
II-3	ARRANGE DIRECTIONS IN LOGICAL ORDER	4	86.58	89.50
IV-1	INCLUDE NEC INFO FOR MESSAGE OR REQUEST	4	85.30	91.75
IV-2	DETERMINE MISSING OR IRRELEVANT INFO	4	85.63	91.25
V-1	PROPER FORMAT OF A BUSINESS LETTER	9	62.48	76.44
V-2	PROPER FORMAT OF A FRIENDLY LETTER	8	70.79	81.13
VI-1	COMPLETE COMMON FORM OR APPLICATION	9	82.64	87.22
VII-1	SPELLING	9	88.30	91.11
VIII-1,2, 3,4	PERIODS, QUEST MARKS, EXCL PTS, COMMAS, APOS	16	78.63	86.44
VIII-5	QUOTATION MARKS	4	66.13	70.75
VIII-6	COLONS	4	64.15	83.75
IX-1	CAPITALIZATION	10	87.59	91.10
	Total =	119	80.00	86.05

Table 8
A Comparison of the Average Difficulty Levels (percentage correct) on on the Initial Pilot and the Live Administration of the AHSGE for Language

Competency Number	Competency Description	# Items	Form 36 Item Pilot	Spring 85 Live Administration
I-1	PRONOUN-ANTECEDENT AGREEMENT	8	93.38	95.25
I-2	CORRECT FORMS OF NOUNS AND VERBS	12	80.17	81.58
II-1	IDEAS IN SEQUENTIAL ORDER	4	80.94	88.25
II-2	SENTENCES ORGANIZED INTO PARAGRAPH	6	80.51	80.17
II-3	ARRANGE DIRECTIONS IN LOGICAL ORDER	4	86.73	86.25
IV-1	INCLUDE NEC INFO FOR MESSAGE OR REQUEST	4	85.28	96.00
IV-2	DETERMINE MISSING OR IRRELEVANT INFO	4	85.55	93.00
V-1	PROPER FORMAT OF A BUSINESS LETTER	6	62.34	77.00
V-2	PROPER FORMAT OF A FRIENDLY LETTER	5	70.59	83.60
VI-1	COMPLETE COMMON FORM OR APPLICATION	8	82.54	93.13
VII-1	SPELLING	8	88.33	90.75
VIII-1,2, 3,4	PERIODS, QUEST MARKS, EXCL PTS, COMMAS, APOS	14	78.62	85.29
VIII-5	QUOTATION MARKS	4	66.15	77.75
VIII-6	COLONS	4	64.10	70.25
IX-1	CAPITALIZATION	9	87.55	89.78
	Total = 100		80.00	86.21

Summary

This paper has provided a comparison of the difficulty level of common items measuring competencies on the initial item pilot and the live administration. This comparison demonstrates the stability of the difficulty level of the common items from the initial item pilot to the live administration. The stability of the difficulty level supports the validation of the test development procedure.

References

Halpin, G. & Steele, D. J. (1984, April). The empirical approach: Method and Results. In J. E. McLean (Chair), A multimethod approach for establishing the passing score on the Alabama High School Graduation Examination. Symposium conducted at the meeting of the National Council on Measurement in Education, New Orleans.

Halpin, G. & Steele, D. J. (1985, April). The Alabama High School Graduation Examination: Test Development. In R. W. Roth (Chair), The Alabama High School Graduation Examination Experience. Syumposium conducted at the meeting of the national Council on Measurement in Education, Chicago.

Hardy, R. (1984, Fall). Measuring instructional validity: A report of an instructional validity for the Alabama High School Graduation Examination. Journal of Educational Measurement, 21, (3), 291-301.

Hess, A. C. (1984, April). Policy implications of the standard setting process. In J. E. McLean (Chair), A multimethod approach for establishing the passing score on the Alabama High School Graduation Examination. Symposium conducted at the meeting of the National Council on Measurement in Education, New Orlenas.

Lockwood, R. E. & McLean, J. E. (1984). Equating of the Alabama High School Graduation Examination. Paper presented in symposium at the meeting of the Mid-South Educational Research Association, New Orleans.

McLean, J. E. (1984). Introduction to the problem and methods. In J. E. McLean (Chair), A multimethod approach for establishing the passing score on the Alabama High School Graduation Examination. Symposium conducted at the meeting of the National Council on Measurement in Education, New Orleans.

Roth, R. W. (1984, April). The judgmental approach: Method and results. In J. E. McLean (Chair), A multimethod approach for establishing the passing score on the Alabama High School Graduation Examination. Symposium conducted at the meeting of the National Council on Measurement in Education, New Orleans.

Steele, D. J. (1984). Practical aspects of developing a high school graduation examination: Technical process of item development--general overview. Paper presented in symposium at the meeting of the Mid-South Educational Research Association, New Orleans.

Teague, W. (1983). Reading, language, mathematics item specifications for the Alabama High School Graduation Examination. Montgomery, AL: Alabama State Department of Education.